

# Northumbria Research Link

Citation: Isupov, Michail N., Boyko, Konstantin M., Sutter, Jan-Moritz, James, Paul, Sayer, Christopher, Schmidt, Marcel, Schönheit, Peter, Nikolaeva, Alena Yu., Stekhanova, Tatiana N., Mardanov, Andrey V., Ravin, Nikolai V., Bezsudnova, Ekaterina Yu., Popov, Vladimir O. and Littlechild, Jennifer A. (2019) Thermostable Branched-Chain Amino Acid Transaminases From the Archaea *Geoglobus acetivorans* and *Archaeoglobus fulgidus*: Biochemical and Structural Characterization. *Frontiers in Bioengineering and Biotechnology*, 7. p. 7. ISSN 2296-4185

Published by: Frontiers

URL: <https://doi.org/10.3389/fbioe.2019.00007>  
<<https://doi.org/10.3389/fbioe.2019.00007>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/43271/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



# Thermostable Branched-Chain Amino Acid Transaminases From the Archaea *Geoglobus acetivorans* and *Archaeoglobus fulgidus*: Biochemical and Structural Characterization

## OPEN ACCESS

### Edited by:

Roberto Contestabile,  
Sapienza University of Rome, Italy

### Reviewed by:

Stefano Bruno,  
University of Parma, Italy  
Samanta Raboni,  
University of Parma, Italy

### \*Correspondence:

Vladimir O. Popov  
vpopov@exchange.inbi.ras.ru  
Jennifer A. Littlechild  
j.a.littlechild@exeter.ac.uk

### Specialty section:

This article was submitted to  
Bioprocess Engineering,  
a section of the journal  
Frontiers in Bioengineering and  
Biotechnology

**Received:** 26 October 2018

**Accepted:** 08 January 2019

**Published:** 24 January 2019

### Citation:

Isupov MN, Boyko KM, Sutter J-M,  
James P, Sayer C, Schmidt M,  
Schönheit P, Nikolaeva AY,  
Stekhanova TN, Mardanov AV,  
Ravin NV, Bezudnova EY, Popov VO  
and Littlechild JA (2019) Thermostable  
Branched-Chain Amino Acid  
Transaminases From the Archaea  
*Geoglobus acetivorans* and  
*Archaeoglobus fulgidus*: Biochemical  
and Structural Characterization.  
Front. Bioeng. Biotechnol. 7:7.  
doi: 10.3389/fbioe.2019.00007

Michail N. Isupov<sup>1</sup>, Konstantin M. Boyko<sup>2</sup>, Jan-Moritz Sutter<sup>3</sup>, Paul James<sup>1</sup>,  
Christopher Sayer<sup>1</sup>, Marcel Schmidt<sup>3</sup>, Peter Schönheit<sup>3</sup>, Alena Yu. Nikolaeva<sup>2</sup>,  
Tatiana N. Stekhanova<sup>2</sup>, Andrey V. Mardanov<sup>2</sup>, Nikolai V. Ravin<sup>2</sup>,  
Ekaterina Yu. Bezudnova<sup>2</sup>, Vladimir O. Popov<sup>2\*</sup> and Jennifer A. Littlechild<sup>1\*</sup>

<sup>1</sup> Henry Wellcome Building for Biocatalysis, Biosciences, University of Exeter, Exeter, United Kingdom, <sup>2</sup> Research Center of  
Biotechnology of the Russian Academy of Sciences, Moscow, Russia, <sup>3</sup> Institut für Allgemeine Mikrobiologie,  
Christian-Albrechts-Universität Kiel, Kiel, Germany

Two new thermophilic branched chain amino acid transaminases have been identified within the genomes of different hyper-thermophilic archaea, *Geoglobus acetivorans*, and *Archaeoglobus fulgidus*. These enzymes belong to the class IV of transaminases as defined by their structural fold. The enzymes have been cloned and over-expressed in *Escherichia coli* and the recombinant enzymes have been characterized both biochemically and structurally. Both enzymes showed high thermostability with optimal temperature for activity at 80 and 85°C, respectively. They retain good activity after exposure to 50% of the organic solvents, ethanol, methanol, DMSO and acetonitrile. The enzymes show a low activity to (*R*)-methylbenzylamine but no activity to (*S*)-methylbenzylamine. Both enzymes have been crystallized and their structures solved in the internal aldimine form, to 1.9 Å resolution for the *Geoglobus* enzyme and 2.0 Å for the *Archaeoglobus* enzyme. Also the *Geoglobus* enzyme structure has been determined in complex with the amino acceptor  $\alpha$ -ketoglutarate and the *Archaeoglobus* enzyme in complex with the inhibitor gabaculine. These two complexes have helped to determine the conformation of the enzymes during enzymatic turnover and have increased understanding of their substrate specificity. A comparison has been made with another (*R*) selective class IV transaminase from the fungus *Nectria haematococca* which was previously studied in complex with gabaculine. The subtle structural differences between these enzymes has provided insight regarding their different substrate specificities.

**Keywords:** thermophilic archaea, branched-chain aminotransferases, substrate specificity, X-ray structural analysis, biocatalysis

## INTRODUCTION

Transaminases (TAMs; aminotransferases; EC2.6.1.-) catalyse a transfer of an amino group between a donor substrate and an acceptor molecule (Mehta et al., 1993). They are involved in the metabolism of most natural amino acids and collectively have a broad substrate range. This makes them very useful for the biocatalytic synthesis of chiral amino acid analogs (Leuchtenberger et al., 2005; Tufvesson et al., 2011) such as *tert*-leucine (Taylor et al., 1998) and homophenylalanine (Cho et al., 2003).

The TAMs belong to a large group of enzymes that utilize pyridoxal 5'-phosphate (PLP) (Hayashi, 1995; Jansonius, 1998). PLP is the biologically active form of vitamin B6 (Braunstein and Shemyakin, 1953; Metzler et al., 1954). The mechanism of these enzymes involves the PLP cofactor forming a covalent bond (a Schiff base) with the NZ atom of the active-site lysine (internal aldimine). In the first half of the TAM reaction, a Schiff base is formed between the donor substrate and PLP (external aldimine). After several reaction stages, the substrate amino group is transferred to the cofactor to form pyridoxamine 5'-phosphate (PMP) with the release of the ketone product. In the second half-reaction the PMP amino group is transferred to an acceptor ketone or aldehyde to form the PLP internal aldimine and a primary amine or a chiral amino product. Most TAMs use an  $\alpha$ -ketoglutarate (AKG) acceptor substrate to form a glutamate product. Some TAMs, however utilize the pyruvate acceptor to form alanine and are usually referred to as substrate:pyruvate TAMs (Ward and Wohlgemuth, 2010; Sayer et al., 2012).

The TAMs are currently widely used in industrial biocatalysis for both the resolution and the asymmetric synthesis of chiral amines which are used as building blocks for drug intermediates. There is specific interest in TAMs capable of the production of both (S) and (R)-chiral intermediates of non  $\alpha$ -amino acid substrates. These enzymes catalyse amination of substrates with a ketone/aldehyde component distal to the carboxyl group [ $\omega$ -TAMs; (Malik et al., 2012)] or with no carboxyl group [amine TAMs; (Shin et al., 2003)]. Several Pfam class III (S)-selective  $\omega$ TAMs (Punta et al., 2012) from *Vibrio fluvialis* (Shin et al., 2003), *Chromobacterium violaceum* and *Pseudomonas aeruginosa* (Kaulmann et al., 2007), *Arthrobacter* sp. and *Bacillus megaterium* (van Oosterwijk et al., 2016) have been biochemically and structurally characterized (Humble et al., 2012; Midelfort et al., 2013; Sayer et al., 2013). They show activity toward the (S)-isoform of  $\alpha$ -methylbenzylamine [(S)-MBA] and also toward other amines and  $\omega$ -amino acids such as  $\beta$ -alanine. The amino acceptor can be pyruvate and a variety of other ketones and aldehydes (Kaulmann et al., 2007).

Most  $\alpha$ -TAMs and (S)- $\omega$ TAMs, are specific toward the (S)-enantiomer of the substrate and have a PLP type I fold. Enzymes of the Pfam TAM class IV have a PLP type IV fold and catalyse transamination of branched chain L-amino acids (BCAAs), D-amino acids or  $\omega$ -amino acids. The catalysis in class IV TAMs occurs on the *re*-side of the cofactor, which differs from most of the other PLP enzymes where the reactions occur on the *si*-side. Since both (S) and (R)-enantiomers can be turned over within the same PLP type IV scaffold, enzymes capable of

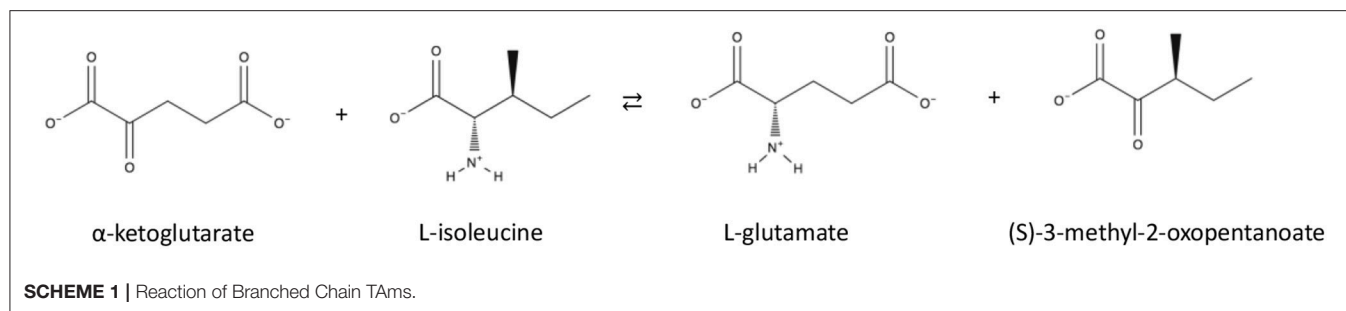
transamination of (R)- $\omega$ -substituted amines are found within this class of TAMs. By comparing the sequence for class IV enzymes which are potentially capable of (R)- $\omega$ TAM activity a sequence fingerprint has been established (Höhne et al., 2010). Biochemical characterization of a number of these enzymes has demonstrated a high enantioselectivity toward the substrates (R)- $\alpha$ -methylbenzylamine [(R)-MBA], (2R)-amino-4-phenylbutane and (2R)-aminoheptane.

The production of the (R)- $\omega$ -substituted chiral amines is a challenging process using conventional chemical synthesis. Therefore, enzymes capable of catalyzing the selective transamination of such compounds have received significant interest. An example of such an enzyme is a mutant (R)-specific transaminase from *Arthrobacter* sp. (ArRMut11), created in a collaboration between Codexis and Merck, which was able to catalyse specific amination of the sterically demanding 1,3-ketoamides to generate the (R)-chiral amine for the manufacture of Sitagliptin, a treatment for type II diabetes (Savile et al., 2010). This mutant ArRMut11 enzyme has been reported in reactions for amination of bicyclic ketones including tetralone (Mutti et al., 2011; Pressnitz et al., 2013). The crystal structures have been reported for several class IV TAM enzymes including the fungal (R)-amine:pyruvate  $\omega$ TAMs from *Aspergillus fumigatus* (Thomsen et al., 2014), *Aspergillus terreus* (Łyskowski et al., 2014) and *Nectria haematococca* (Sayer et al., 2014), and the bacterial (R)-selective  $\omega$ TAM ArRMut11 from *Arthrobacter* sp (Guan et al., 2015).

There is an increasing demand for enzymes which are more robust to the demanding conditions used in industry. Enzymes found in thermophilic organisms have increased thermostability and are more tolerant to organic solvents and proteolytic cleavage. Solvent stability is advantageous since non-natural substrates used industrially often require the addition of organic solvents to the reaction mixture for substrate solubilisation (Littlechild et al., 2007). Also the biocatalytic process can be carried out at elevated temperatures where many non-natural substrates have improved solubility when using a thermostable enzyme which can be reused through several reaction cycles. This reduces the overall cost of the enzyme in the industrial process which is often a limitation in the development of a biocatalytic process.

Branched chain TAMs (BCATs) catalyse reversible transamination of branched chain amino acids (shown in **Scheme 1**). Recently, archaeal thermophilic BCATs have been biochemically characterized from *Thermococcus* sp (Uchida et al., 2014) and biochemically and structurally studied from the thermophile *Thermoproteus uzoniensis* (Boyko et al., 2016).

This paper reports the identification, biochemical and structural characterization of two new thermostable archaeal class IV TAMs from *Geoglobus acetivorans* (Querellou et al., 2009) and *Archaeoglobus fulgidus* (Stetter, 1988). Both of these hyperthermophilic archaea have been isolated from different deep sea hydrothermal vents and they share 79.2% sequence identity. The structures of the *G. acetivorans* enzyme have been determined in the internal aldimine form and in complex with the amino acceptor AKG and the structures of the *A. fulgidus* enzyme in the internal aldimine form and in complex with the inhibitor



gabaculine. The different structural complexes of these related enzymes have given further insight into the overall mechanism of BCATs and their high stability for industrial application and their substrate specificity.

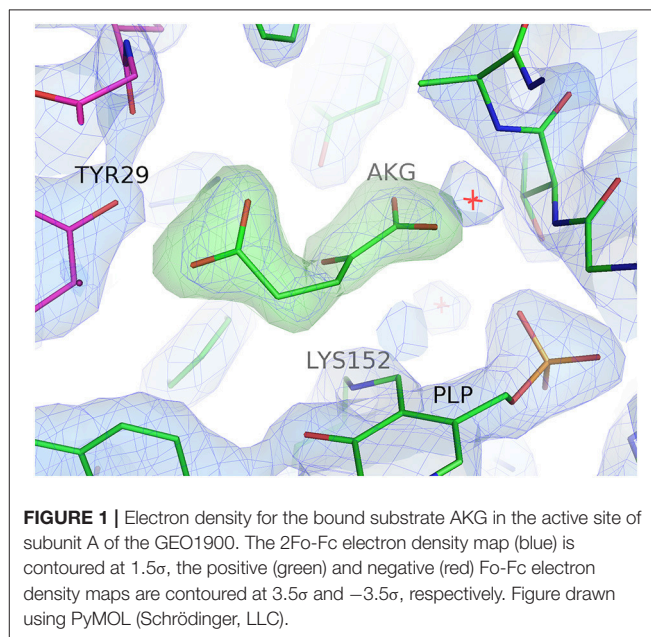
## RESULTS AND DISCUSSION

### Enzyme Cloning, Expression, and Purification

The genes encoding two putative BCATs were identified in the genomes of *G. acetivorans* (Mardanov et al., 2011) and *A. fulgidus* (Klenk et al., 1997). Both proteins called GEO1900 and AF0933 have been cloned and over-expressed in a soluble form in *Escherichia coli* and have been purified to homogeneity using metal affinity and size exclusion chromatography. The recombinant BCATs GEO1900 (MW of subunit 32.6 kDa, 292 amino acids) and AF0933 (MW of subunit 32.4 kDa, 290 amino acids) are closely related with a sequence identity of 79.2% and 94.8% similarity. When purified by high resolution gel filtration chromatography the native molecular weight of the two enzymes varied with the GEO1900 approximately 70 kDa, indicating that the enzyme was a homodimer with small amount of tetramer in solution (**Figure S1**). However, the AF0933 enzyme had a native molecular mass of approximately 220 kDa as determined by size exclusion chromatography, indicating that it forms a homo-hexamer with only small amounts of a homo-dimeric enzyme observed (**Figure S2**). Both proteins showed absorption at 420 nm indicating that the cofactor PLP was bound in the aldimine form (data not shown). Both of the GEO1900 and AF0933 proteins run as a single band of 36 kDa on SDS-PAGE which is in line with the calculated molecular mass of the His-tagged proteins (35.4 kDa).

### Biochemical Characterization of GEO1900 and AF0933

Both enzymes, GEO1900 and AF0933, showed significant activity toward keto analogs of BCAAs (**Tables 1, 2**). The apparent kinetic parameters for the AF0933 catalyzed transamination reaction between L-glutamate and the branched chain oxoacids were similar with a preference for 3-methyl-2-oxopentanoate. The extrapolated  $V_{\max}$  and  $K_m$  values for L-glutamate were significantly higher (**Table 2** and **Figure S7**). These values of  $V_{\max}$  and  $K_m$  point to the low affinity binding of L-Glu in the active site. This was observed earlier for other BCATs from *P.*



*aeruginosa* and *Gluconobacter oxydans* (Norton and Sokatch, 1970; Tachiki and Tochikura, 1973; Kanda et al., 1995). For GEO1900 the extrapolated  $V_{\max}$  and  $K_m$  values for pyruvate and AKG were ( $1.9 \pm 0.1$  U/mg and  $10.6 \pm 1.9$  mM) and ( $18.2 \pm 1.2$  U/mg and  $1.9 \pm 0.6$  mM), respectively. The calculated catalytic efficiency constant for GEO1900 toward AKG was found to be higher than toward pyruvate,  $9.8 \text{ s}^{-1} \text{ mM}^{-1}$  compared to  $0.1 \text{ s}^{-1} \text{ mM}^{-1}$ . Both values fall within the range of the calculated values for the canonical BCATs (Bezsudnova et al., 2017).

As for the activity toward amino donors the profile of activity toward amino acids for GEO1900 (**Table S1**) was similar to that reported for related BCATs from *E. coli*, *Bacillus brevis* and *Pseudomonas aeruginosa* (Norton and Sokatch, 1970; Lee-Peng et al., 1979; Inoue et al., 1988; Kanda et al., 1995; Yu et al., 2014). GEO1900 was highly active toward branched-chain amino acids and their isomers. The AF0933 catalyzed the transamination between L-valine and AKG with a specific activity of 5.76 U/mg and between L-alanine and 3-methyl-2-oxobutanoic acid with a lower specific activity of 0.16 U/mg. When the enantio-preference of the GEO1900 and AF0933 in the reaction with the (R) and (S)-isomers of methylbenzylamine



**TABLE 1 |** The specific activity of GEO1900 and AF0933 toward different keto acids.

Amino acceptor	Specific activity (U/mg)	
	GEO1900 <sup>a</sup>	AF0933 <sup>b</sup>
3-methyl-2-oxobutyrate	n.a.	5.14
4-methyl-2-oxopentanoate	7.7	3.05
3-methyl-2-oxopentanoate	7.1	3.23
2-oxopentanoate	n.a.	6.51
2-oxohexanoate	n.a.	4.26
2-oxooctanoate	n.a.	2.4
2-oxobutyrate	1.3	n.a.
pyruvate	0.56	0.73
phenylglyoxylate	n.a.	4.55
$\beta$ -phenylpyruvate	n.a.	2.84
indole-3-pyruvate	n.a.	1.02

<sup>a</sup> 5 mM L-glutamate was used as the amino donor and 5 mM keto acid was used at 65°C in 50 mM phosphate buffer, pH 8.0; supplemented with 60  $\mu$ M PLP.

<sup>b</sup> 10 mM L-glutamate was used as the amino donor and 10 mM keto acid was used at 70°C in 100 mM Tris-HCl buffer pH 7.5 supplemented with 0.1 mM PLP.

Not all donor acceptor pairs were tested for both enzymes.

**TABLE 2 |** The apparent kinetic parameters of the transamination reaction catalyzed by AF0933.

Substrate	V <sub>max</sub> (U/mg)	K <sub>m</sub> (mM)	k <sub>cat</sub> /K <sub>m</sub> (s <sup>-1</sup> mM <sup>-1</sup> )
3-methyl-2-oxobutyrate <sup>a</sup>	6.3	0.42	8.85
4-methyl-2-oxopentanoate <sup>a</sup>	5.5	0.25	12.98
3-methyl-2-oxopentanoate <sup>a</sup>	3.9	0.13	17.7
L-glutamate <sup>b</sup>	64.6*	121	0.32

<sup>a</sup> At 10 mM L-glutamate.

<sup>b</sup> At 5 mM 3-methyl-2-oxobutyrate.

\* The V<sub>max</sub> for L-glutamate is an extrapolated value calculated at saturated concentration of co-substrate.

(MBA) was investigated using the acetophenone assay they both showed a small but significant activity toward (R)-MBA at pH 7.5, with a value of 2 and 10 mU/mg with 2-oxobutyrate as the keto substrate, respectively. Neither BCAT showed any activity toward (S)-MBA.

The temperature optimum for the transamination reaction catalyzed by GEO1900 and AF0933 was found to be 80 and 85°C, respectively (**Figures S3, S5**). The pH optimum for amination of keto acids with L-Glu was between pH 7.0 and 8.0 for both enzymes. The thermostability of GEO1900 and AF0933 was evaluated using the analysis of the residual activity after pre-incubation of the enzyme for 2 h at different temperatures. Significant loss of activity of GEO1900 was observed at temperatures higher than 50°C. The pre-incubation of the enzyme at 70°C resulted in 25% loss of activity after 2 h (**Figure S4**). At the same time more than 60% of the AF0933 residual activity remained after pre-incubation at 75°C for 2 h (**Figure S6**). The pre-incubation of AF0933 at 85°C for 60 min resulted in a drop of the residual activity to 30%.

The solvent stability of GEO1900 and AF0933 was tested in ethanol, methanol, DMSO and acetonitrile. After 1 h of incubation in up to 50% of these solvents GEO1900 retained close to 100% activity at both 20°C and 55°C. For AF0933 after 1 h incubation of the enzyme in up to 50% of the tested solvents at the similar temperatures 80% of the original activity remained.

## Structural Studies

The GEO1900 holoform (internal aldimine form; GEO1900\_holo) and the unproductive complex of holoenzyme with the keto substrate AKG-(GEO1900\_AKG) have been determined to 1.9 and 2.2 Å resolution, respectively, in the trigonal space group P3<sub>2</sub>21. There are three GEO1900 subunits per asymmetric unit, which form a tight hexamer with 32 point group symmetry with their symmetry mates. The AF0933 structure has been determined in the holoform (AF0933\_holo) and in complex with the inhibitor gabaculine (GABC) (AF0933\_GABC) to 2.1 and 2.0 Å resolution, respectively, in the orthorhombic P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> space group. The asymmetric unit contains a hexameric AF0933 molecule which is very similar to GEO1900 hexamer observed in the crystal. All of the structures have been refined to low R-factors with good stereochemical parameters as shown in **Table 3**. In the GEO1900 structures two N-terminal amino acids which are not present in AF0933, were not defined in the electron density. The interdomain loop residues 121–129 could be fully built into the subunits A and B in GEO1900\_AKG structure. The bound ligand AKG was found only in subunit A and was clearly defined in the electron density (**Figure 1**). The interdomain loop could only be fully built in one subunit (D) out of the six subunits in both the AF0933\_holo and AF0933\_GABC structures. The crystal contacts in AF0933 appear to be contributing to the lower disorder of the loop in the D subunit. The inhibitor, gabaculine was not present in the electron density in the subunit F of AF0933, therefore this subunit was built as an internal aldimine in the AF0933\_GABC structure. In other subunits the gabaculine-PLP adduct was modeled with partial occupancy in the range of 0.6–0.7. There are no residues in a *cis* conformation in the well-defined regions of the GEO1900 and AF0933 structures. Residue Glu249 is a Ramachandran plot outlier (Ramakrishnan and Ramachandran, 1965) in all subunits of both AF0933 structures. The corresponding Glu251 has similar main chain angles in the GEO1900 structures, but lies in a generously allowed region of the Ramachandran plot (Laskowski et al., 1993). The fungal class IV Tams from *Aspergillus* and *Nectria* species (Łyskowski et al., 2014; Sayer et al., 2014; Thomsen et al., 2014) have a Gly residue at this position. Many residues in both of the GEO1900 and AF0933 structures have been modeled with multiple conformations of some amino acid side chains. The residues Gln107 and Asn108 were modeled with alternative conformations of the main chain in most subunits of the AF0933\_holo structure and in the subunits E and F of the AF0933\_GABC structure.

## Subunit Structure

The subunit of the archaeal BCATs has a PLP type IV fold typical of other BCATs, D-amino acid Tams and (R)-selective  $\omega$ Tams

**TABLE 3 |** Summary of data processing and refinement statistics.

	Holo enzyme GEO1900	$\alpha$ -ketoglutarate complex GEO1900	Holo enzyme AF0933	Gabaculine complex AF0933
Diffraction source	BL41XU, Spring8	BL41XU, Spring8	I04-1, Diamond	I04-1, Diamond
Wavelength (Å)	1.0	1.0	0.92	0.92
Space group	P3 <sub>2</sub> 21	P3 <sub>2</sub> 21	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
a, b, c (Å)	117.65, 117.65, 135.98	117.31, 117.31, 135.31	72.92, 139.12, 167.57	73.21, 140.23, 168.32
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 120	90, 90, 120	90, 90, 90	90, 90, 90
Resolution range (Å)	58.83–1.90 (2.00–1.90) <sup>a</sup>	67.66–2.20 (2.26–2.20)	69.60–2.10 (2.14–2.10)	67.17–1.98 (2.01–1.98)
Number of unique reflections	85638 (11949)	55018 (4457)	99985 (4906)	119864 (5675)
Completeness (%)	99.6 (98.6)	99.9 (99.8)	99.9 (99.6)	99.0 (95.6)
Average redundancy	19.8 (19.3)	17.7 (18.4)	6.3 (5.4)	5.6 (4.7)
$\langle I/\sigma(I) \rangle$	31.75 (4.0)	15.1 (4.9)	10.0 (0.8)	11.2 (0.8)
R <sub>meas</sub> (%) <sup>b</sup>	6.8 (79.2)	13.3 (67.4)	14.7 (251.1)	12.8 (233.5)
CC <sub>1/2</sub> (%) (Diederichs and Karplus, 1997)	100 (92.9)	99.8 (92.6)	99.7 (29.2)	99.7 (24.9)
Overall B factor from Wilson plot (Å <sup>2</sup> ) <sup>c</sup>	39.8	42.3	51.3	44.8
R <sub>fact</sub> (%)	15.3	15.0	18.3	18.1
R <sub>free</sub> (%)	18.8	20.8	21.8	20.9
Refined protein atoms	6671	6836	14432	14440
Refined solvent atoms	626	508	754	828
Refined ligand atoms	0	10	0	125
<b>AVERAGE B FACTOR (Å<sup>2</sup>)</b>				
Protein	37.9	38.2	45.8	42.7
Solvent	45.3	40.9	50.6	47.9
R.m.s.d. bond lengths (Å)	0.021	0.019	0.013	0.013
R.m.s.d. bond angles (°)	1.99	2.11	1.78	1.69
<b>RAMACHANDRAN PLOT ANALYSIS, RESIDUES IN (%)<sup>d</sup></b>				
Most favored regions	90.5	89.9	90.6	90.7
Additional allowed regions	8.7	8.8	8.6	8.5
Generously allowed regions	0.8	1.1	0.4	0.4
Disallowed regions	0	0.3	0.4	0.4

<sup>a</sup>Values for the highest resolution shell are given in parentheses.

<sup>b</sup> $R_{meas} = \sum_h [m/(m-1)]^{1/2} \sum_i |I_{h,i} - \langle I_h \rangle| / \sum_h \sum_i I_{h,i}$  (Karplus and Diederichs, 2012).

<sup>c</sup>Wilson B-factor was estimated by SFCHECK (Vaguine et al., 1999).

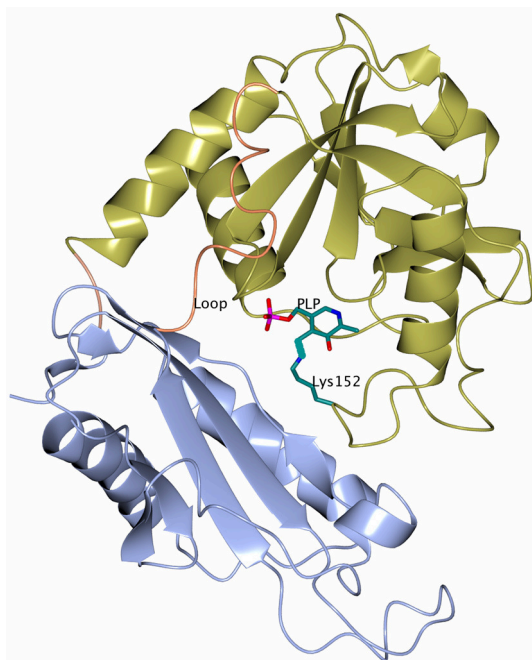
<sup>d</sup>Ramachandran plot analysis was performed by PROCHECK (Laskowski et al., 1993).

(Mehta et al., 1993; Höhne et al., 2010) and consists of two  $\alpha/\beta$  domains, a small domain and a large domain connected by two inter-domain loops (Peisach et al., 1998; Okada et al., 2001; Goto et al., 2003; Hirotsu et al., 2005; Boyko et al., 2016) (Figure 2).

The small domain (residues 1–119 and 287–292 in GEO1900) is composed of three  $\alpha$ -helices wrapped around a seven stranded antiparallel  $\beta$ -sheet with topology  $-1, +4x, +1, -2x, -1, +4x$  as defined by PROMOTIF (Richardson, 1981; Kabsch and Sander, 1983; Hutchinson and Thornton, 1996) with the C-terminus contributing to the last  $\beta$ -strand. The large PLP-binding domain (residues 130–285) has six  $\alpha$ -helices surrounding an eight stranded  $\beta$ -sheet with topology  $-1x, -1, +5, +1, +1, -3, -1$  and direction  $- + + + - + -$ . Both domains are connected via a short link at C-terminus (residues 286–287) and a long inter-domain loop (residues 120–129).

## Quaternary Structure of BCATs

A dimer formation is required for the BCAT catalytic activity, since residues from both subunits in the dimer contribute to the active site (Figure 3). However, both GEO1900 and AF0933 form hexamers (trimers of dimers) in the crystal (Figure 4), with the GEO1900 hexamer located on a crystallographic dyad. Dimerization of GEO1900 results in 2400 Å<sup>2</sup> or 18% of each subunit solvent accessible surface area being buried (Krissinel and Henrick, 2007). Hexamer formation of GEO1900 in the crystal leads to an additional 14% of each subunit surface area being buried. Similarly, dimerization of AF0933 buries 19% of subunit surface area with additional 14% buried upon hexamer formation. This is very much in line with other bacterial BCATs that form a hexameric structure in the crystal (Okada et al., 2001; Chen et al., 2012; Boyko et al., 2016), with organization of all hexamers being very conserved, the core rmsd in C $\alpha$  positions are comparable for both dimers and hexamers in pairwise alignment.



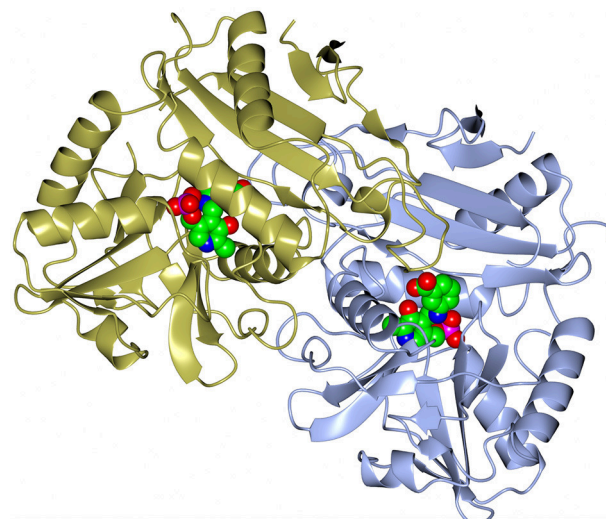
**FIGURE 2** | A cartoon representation of the subunit of GEO1900 with the large domain shown in gold and the small domain shown in ice blue. The loops connecting the domains are shown in coral. The cofactor PLP and the active site lysine 152 are shown as stick models with carbon atoms shown in green. **Figures 2–5, 7, 8** are drawn using CCP4mg (McNicholas et al., 2011).

Interestingly, known structures of other members of group IV proteins—D-amino acid Tams, bacterial and fungal amine Tams are homodimers (Peisach et al., 1998; Sayer et al., 2014; Thomsen et al., 2014). In this context a dimeric structure in solution observed by size exclusion chromatography for GEO1900 and *T. uzoniensis* BCAT (Boyko et al., 2016), and a small dimeric fraction for AF0933 suggests an equilibrium between hexameric and dimeric forms, since only dimer is required for catalysis. The AF0933 is predominantly hexameric in solution (**Figure S2**) which may account for its increased thermal optimum over that of GEO1900.

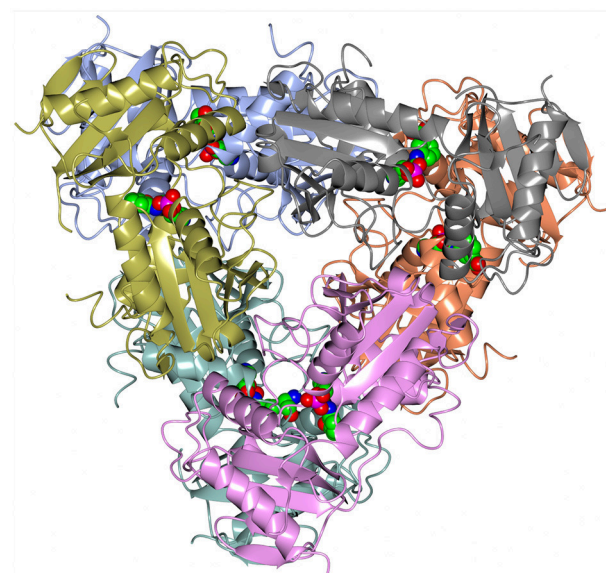
## Active Site

The active site of the archaeal BCATs is located at the bottom of the cleft formed by residues of both domains of one subunit as well as of the small domain of an adjacent subunit within the catalytic dimer. The same arrangement has also been observed in other BCATs (Okada et al., 2001) and generally in all enzymes with a PLP type IV fold (Peisach et al., 1998; Sayer et al., 2014; Thomsen et al., 2014). A PLP molecule is covalently bound to the catalytic Lys152 in the GEO1900 active site (Lys150 in AF0933). Since most active site residues are conserved between the two BCATs the GEO1900 numbering will be used henceforth.

In the holo and complex structures of the archaeal BCATs the PLP moiety is clearly visible in the electron density map and has B factors similar to those of the neighboring protein residues, indicating a full occupancy of the cofactor. The PLP



**FIGURE 3** | A cartoon representation of the dimer of the AF0933 BCAT gabaculine complex. The gabaculine-PLP covalent adduct molecules are shown as spheres.



**FIGURE 4** | A cartoon representation of the AF0933 BCAT holo structure in the native hexameric form. The PLP cofactor molecules are shown as spheres.

forms a number of contacts with the protein molecule as found in other BCATs (Hutson, 2001; Okada et al., 2001; Yennawar et al., 2001). Apart from the Schiff base formed by the NZ of Lys152 residue, the pyridine ring of the cofactor forms hydrogen bonds to the conserved Tyr156 and Glu185 residues. The PLP phosphate group forms a salt bridge with Arg53 and a number of hydrogen bonds to the main chain nitrogens and side chains of residues Thr212 and Thr248, and the main chain nitrogen of Ile211, and two water molecules (**Figure S8**). The active site of BCAT can be subdivided into two pockets, A and B.



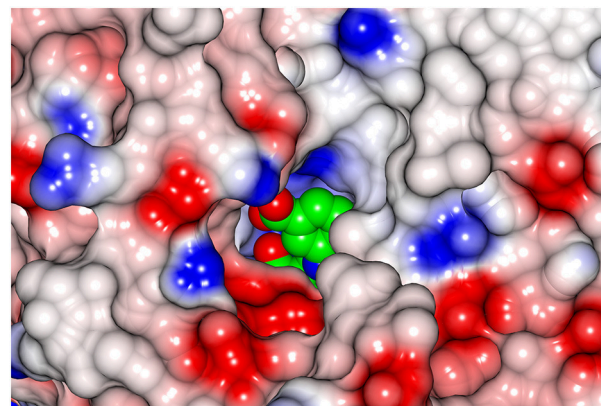
In GEO1900, the B-pocket is formed by the residues Gly36, Arg38, Tyr89, Gly188, Thr248, and Ala249 (**Figure S8**). When compared to the structure of *E.coli* BCAT (Okada et al., 2001) these residues are found to be involved in the binding of the substrate carboxyl group. The A-pocket is lined by conserved residues Arg91, Tyr29', Phe34, Trp120, and Tyr124 as well as the non-conserved residues Leu101' and Leu103' (where ' denotes residues of another subunit).

## Ligand Binding and Enzyme Negative Cooperativity

A bound AKG molecule is clearly visible in the electron density (**Figure 1**) of one subunit of the GEO1900\_AKG complex and has a B-factor comparable to the average for the protein. A contact analysis reveals that  $\alpha$ -carboxylic as well as  $\alpha$ -keto groups of the ligand are tightly bound in the active site via hydrogen bonds with the highly conserved residues Tyr89, Ala249, Thr248, Arg91 as well as Ala250 and Gly188 (via a conserved water molecule D238) (**Figure S8**). Although most of the residues around the  $\gamma$ -carboxylic group of the ligand are hydrophobic, this group is also connected to the protein via several hydrogen bonds with Arg91, Tyr29' and Leu103' (via the water molecule-D923).

Comparative analysis of the holo form of GEO1900 and its complex with AKG reveals that the binding of the acceptor substrate does not result in any significant conformational changes in the active site, except for the rotation of the side chains of Leu101' and Leu103', accompanied by a small main chain displacement (distance between corresponding C $\alpha$ -atoms is about 0.6 Å). In addition the AKG binding residues 120–124 of the inter-domain loop shift toward the active site. This positions the side chains of residues Trp120 and Leu123 so that they shield the ligand from the solvent (**Figures S9, S10**). Interestingly, in the case of the two archaeal BCATs the residues Leu123 and Tyr124 are not part of the consensus sequence X-G-X-Y-L (where the first X is an aromatic residue) which is typical of other BCATs, rather this motif is X-G-X-L-Y in both GEO1900 and AF0933. This permutation (Tyr to Leu and *vice versa*) leads to a change in the shape of the active site cavity, making the A-pocket more extended in the direction of Leu123.

The inhibitor gabaculine makes an irreversible complex with the cofactor PLP due to the migration of the double bond of the external aldimine onto the gabaculine ring (**Figure 5**). However, the PLP-gabaculine adduct in the AF0933 complex structure has only partial occupancy of the gabaculine in the enzyme active site. Occupancy of the ligand in subunit F is below the level for an easy build into the model structure, therefore the internal aldimine was modeled into this subunit. The active site of the AF0933 BCAT is much more restrictive for binding of the gabaculine moiety than that of the fungal (*R*)-selective  $\omega$ Tam from *Nectria haematococca* previously studied by us where it was found at 100% occupancy (Sayer et al., 2014). However, when 100:1 excess of gabaculine to PLP was added to the crystallization droplet it was expected to find full occupancy of the PLP-ligand adduct, unless AF0933 BCAT exhibits some negative cooperativity for substrate binding. Binding of the gabaculine inhibitor to one BCAT subunit appears to be sensed by neighboring subunits,



**FIGURE 5** | A view from outside onto the active site cavity entrance using a surface representation of the AF0933 BCAT complex with the inhibitor gabaculine (shown as spheres).

with conformational adjustments preventing the binding of the inhibitor in the other subunits. Negative cooperativity has been reported earlier for other PLP enzymes (Stetefeld et al., 2006; Ruzicka and Frey, 2010) and is in agreement with the fact that the AKG binding in the GEO1900 complex structure is only found in one of the three subunits.

## Inter-domain Loop Conformation

The conformation of the loop 121–129 is significantly different between the BCAT enzymes and other class IV enzymes such as (*R*)-selective *Nectria* Tam where this loop covers the active site, creating an enclosed environment which would facilitate the activation of the amine nitrogen. In the AF0933 BCAT enzyme this loop is poorly ordered and moves away from the active site. In other BCATs including the GEO1900 BCAT this loop is closer to the active site but does not cover it as tightly as seen in the *Nectria* Tam. However, the loop appears to become more ordered in the GEO1900\_AKG complex.

## Comparison of Structures of the Native Enzymes and the Complexes of GEO1900 and AF0933

From biochemical analysis it is known that AKG does not inhibit the activity of GEO1900 in the concentration range tested (**Figure S11**). It would appear that the crystal contacts play a role in increasing the affinity of GEO1900 toward AKG in one single subunit of the enzyme. Superposition of the subunits from the asymmetric unit gave a RMSD between the C $\alpha$  atoms of about 0.2 and 0.5 Å in case of GEO1900\_holo and GEO1900\_AKG, respectively. In the case of the GEO1900\_holo structure the conformations of the flexible loops have small differences between the different subunits. However, these are likely to be due to crystal packing effects. The structure of the complex GEO1900\_AKG demonstrates that a major difference between the subunits can be assigned to a movement of the inter-domain loop which has a sufficiently different conformation in subunit A (as shown in **Figure S9**), where the ligand was bound, when compared to the other subunits. This loop was reported to



shield the active site from the solvent upon the substrate binding (Hirotsu et al., 2005). There are two further differences between the GEO1900\_AKG and GEO1900\_holo structures. The first is the conformation of the loop 264–266, which differs in the A subunit when compared to subunits B and C, with a distance between the matching C $\alpha$ -atoms of 1.0 Å. The second difference is in the region of residues 249–252, where the distance between corresponding C $\alpha$ -atoms of the different subunits reaches 0.6 Å. As the residue displacements were only observed for the A subunit containing bound AKG, it is suggested that this could be the result of main-chain movement due to the substrate binding.

Superposition of the GEO1900\_holo subunits on the GEO1900\_AKG subunits reveals a further major difference in the orientation of the inter-domain loop. This loop is usually disordered in the absence of a bound substrate (Okada et al., 2001; Chen et al., 2012) with exception of the structure of the BCAT from the archaeon *T. uzoniensis* (PDB ID 5CE8; Boyko et al., 2016). In the GEO1900\_holo structure the inter-domain loop is not seen in the electron density for subunits A and C, but in subunit B only two residues 120–121 are not observed (Figure S9). The other residues of the inter-domain loop have higher B factors than the average B factor of the model. It is proposed that the changes in the conformation of the inter-domain loops in GEO1900\_AKG reflect the structuring of this loop during substrate binding. Thus, in the A subunit this loop adopts a conformation which shields the bound substrate from the solvent. However, in the B subunit of the complex the conformation of inter-domain loop is similar to that found in subunit B of the holo\_GEO1900 structure. In the C subunit this loop is not seen in the electron density.

Although the AF0933 BCAT GABC complex was obtained by co-crystallization with gabaculine, differences in the main chain atom positions between the holoenzyme structure and the complex are minimal. No movement of the interdomain loop upon formation of the PLP-gabaculine complex was observed. The carboxyl group of the gabaculine-PLP adduct points away from both distal and proximal carboxyl group binding pockets of the BCAT. This is probably due to the orientation of this adduct being sterically hindered by the shape of the AF0933 BCAT active site cleft.

## Comparison to Structures of Other BCATs and Other CLASS IV Transaminases

The BLAST search found several BCATs with known three-dimensional structure which have a moderate primary sequence similarity to GEO1900 and AF0933. The structure-based sequence alignment (Figure 6), shows the comparison of the GEO1900 and AF0933 sequences with other BCATs and another class IV TAm enzyme from the fungal species *Nectria*. This shows that the *Nectria* enzyme has an extended N-terminus and also a difference in the loop residues 99–108 (AF0933 numbering) between  $\beta$ 5 and  $\beta$ 6 of the N-terminal small domain. This loop contributes to the active site and is much larger in the GEO1900 and AF0933 BCATs (Figure 7).

The superposition of the GEO1900\_holo subunit and the subunits of the AF0933 and other homologous enzymes using

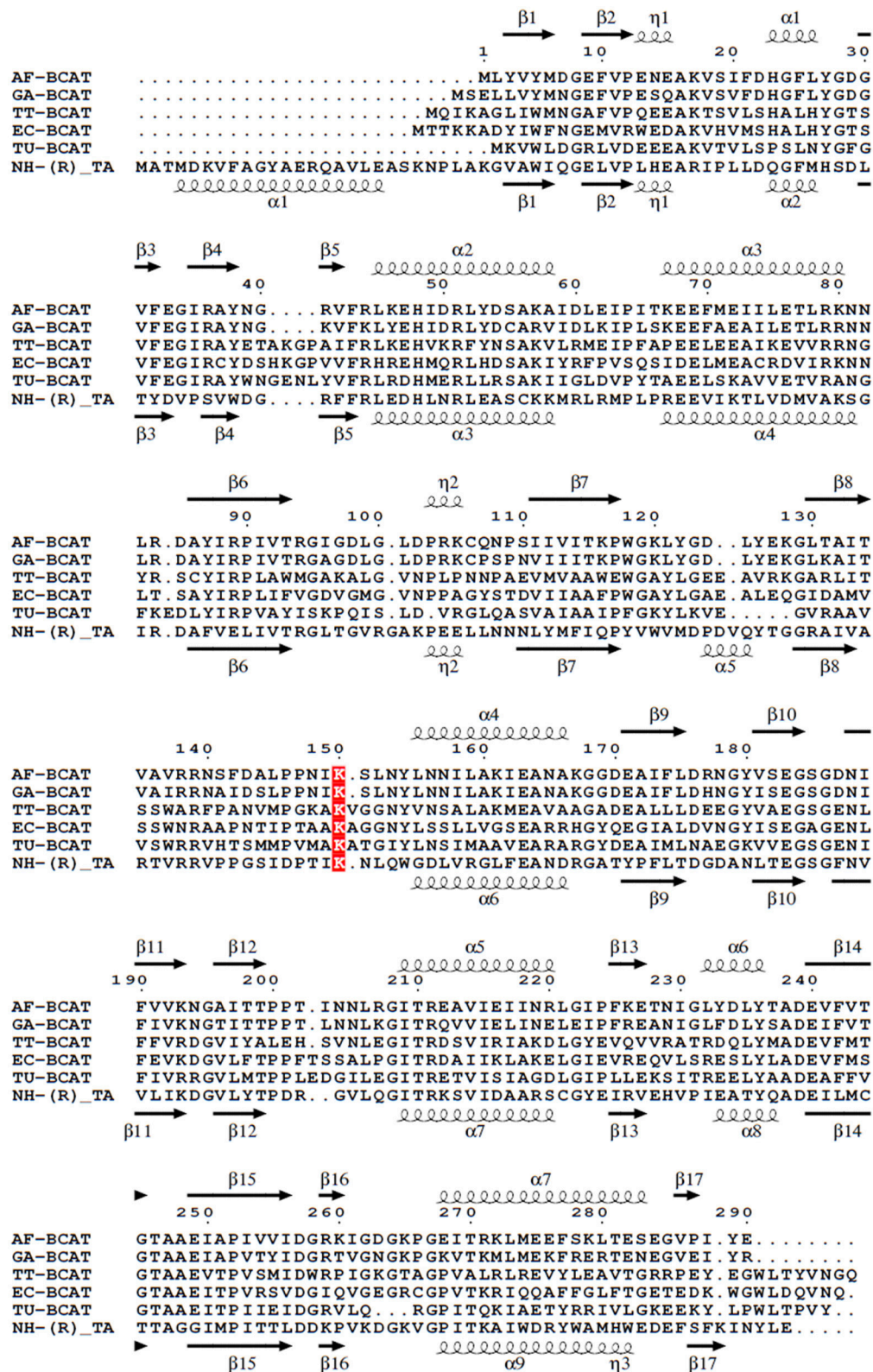
PDBFOLD showed that the fold of the new thermostable BCATs is very similar to related enzymes. Small differences are observed within the flexible loops that connect the secondary structure elements. However, the inter-domain loop conformation is similar between GEO1900\_AKG and the complexes of the *E. coli* BCAT with substrates bound (PDB codes—1I1M and 1I1L).

A comparison of the structures of GEO1900 and AF0933 has been made with other class IV TAmS. The overall structures of these enzymes are similar for DATAs, BCATs and (R)-selective amine:pyruvate TAmS. It appears that the differences are in the amino acid residues lining the active site cavity that determine substrate specificities. The BCAT catalyzed transamination between branched chain L-amino acids and the AKG acceptor produces the branched chain ketoacid and L-glutamate. The DATAs catalyze the exchange of the amino group between D-amino acids and the AKG. The (R)-selective amine:pyruvate TAmS catalyze transfer of an amino group between a (R)- $\omega$ -substrate and pyruvate. Höhne et al. (2010) have identified two signature motifs which distinguish the BCAT and DATA enzymes from other (R)-specific amine:pyruvate TAmS such as the *Nectria* TAm (Sayer et al., 2014). The conserved *Nectria* TAm residues His53, Tyr58, and Val60 belong to the first signature motif and Phe113 to the second motif. These residues contribute to the surfaces of the A and B substrate binding pockets of the active sites to enforce specific binding of the (R)-amine. The residues Ser62 from the first signature motif and Glu115 from the second motif, in the *Nectria* TAm prevent the BCAT activity for this enzyme. In the BCATs the equivalent residues Arg38 and Arg91 (GEO1900 numbering) are responsible for binding the  $\alpha$ -carboxyl and the  $\gamma$ -carboxyl groups of the substrate.

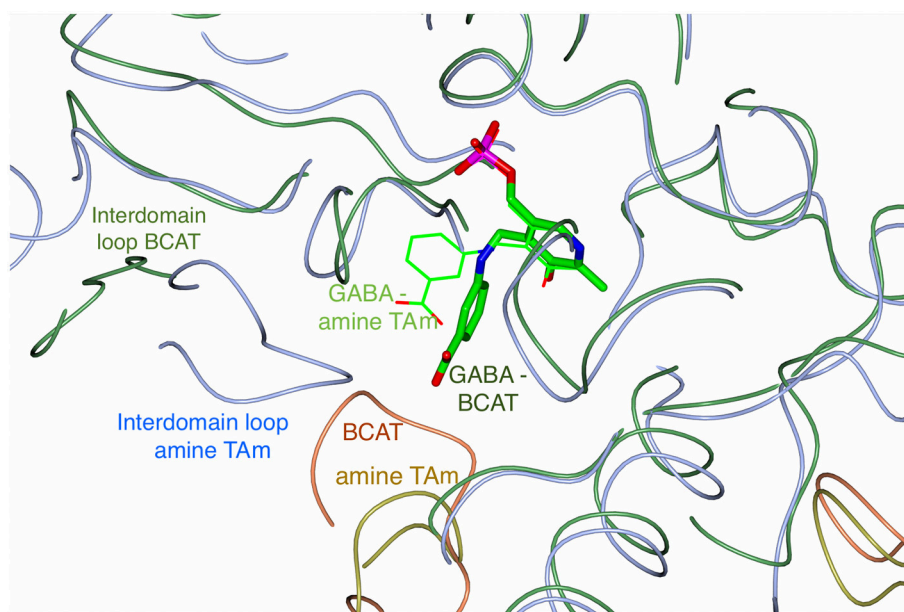
## Structural Basis for GEO1900 and AF0933 Thermostability

Both BCATs GEO1900 and AF0933 demonstrate high thermal and solvent stability. The closest sequence homologs of these enzymes with a known 3-D structure are from another thermostable archaeal BCAT from *T. uzoniensis* (45% sequence identity to AF0933; Boyko et al., 2016), *Thermus thermophilus* (42% identity; PDB ID: 2E1Y) and a mesophilic *E. coli* BCAT (40% identity; Okada et al., 2001). Thermostability of the *E. coli* BCAT does not appear to have been reported, however one can assume it is significantly lower than that of its counterparts from thermophilic organisms, although in some cases such presumption may be erroneous. For instance, the *E. coli* tryptophanase has high thermostability and was purified using a heat treatment step (Dementieva et al., 1994).

Features that are known to confer increased thermostability of proteins include an increase in ionic interactions or salt bridges which offer more thermostability when found in clusters, increased hydrophobicity and the shortening of surface loops (Littlechild et al., 2013). The AF0933 TAm subunit forms up to 27 ion pairs, with five 3-amino acid salt bridge clusters formed. There are no ion pairs between the subunits in the catalytic dimer, however three ion pairs per subunit are formed with neighboring subunits within the hexamer. Residues with multiple side chain conformations in many cases form ion pairs with



**FIGURE 6 |** Multiple sequence alignment of different BCATs, *A. flugidus*, *G. acetivorans*, *T. thermophilus*, *E. coli*, *T. uzoniensis*, *N. haematococca*. Arrows indicate  $\beta$ -strands, and helical curves denote  $\alpha$ -helices of the structure of AF0933 above and *Nectria* TAm below. The active site lysine is highlighted in red. The figure was prepared with ESPript3 (Robert and Gouet, 2014).



**FIGURE 7 |** An overlay of the C $\alpha$  trace of the AF0933 BCAT (green and coral coil) gabaculine complex with the *Nectria* TAM (ice blue and gold coil) gabaculine complex (Sayer et al., 2014) to illustrate the differences between the ligand and loop conformations within the active site between the two enzymes. The PLP gabaculine adduct is shown as a stick model for the AF0933 BCAT structure and as thick lines for the *Nectria* amine TAM structure. Interdomain loops and the loop connecting  $\beta 5$  and  $\beta 6$  of the N-terminal small domain of the adjacent subunit in the catalytic dimer covering the active site cavity, have different conformation between the AF0933 BCAT and the *Nectria* TAM as shown.

different residues for either conformation. The *E. coli* BCAT has 15 ion pairs altogether, including those involved in four 3-amino acid clusters. None of these are with neighboring subunits. Most of salt bridges and other structural features observed in the AF0933 are conserved in GEO1900, therefore comparison here is limited to the former protein.

The AF0933 BCAT has much shorter loops connecting secondary structure elements in comparison with the *E. coli* BCAT. Particularly shortened are the loops connecting strands  $\beta 3$  to  $\beta 4$  and  $\beta 5$  to  $\beta 6$  of the N-terminal domain and  $\beta 5$  to  $\beta 6$  of the large domain. The C-terminal inter-domain loop, which is only 2 residues long in the AF0933 BCAT, is 9 residues long in the *E. coli* enzyme. Shorter loops also mean that there is a higher percentage of residues belonging to the secondary structure, 55.9% in AF0933 BCAT compared to 49.2% in the *E. coli* BCAT. The inter-subunit interface in the catalytic dimer is mainly hydrophobic both in the AF0933 and in the *E. coli* BCAT, therefore it appears that hydrophobic interactions contribute equally to the stability of both proteins.

Thus, an increased number of ion pairs per subunit alongside shorter surface loops are the main contributing factors to the increased thermostability of AF0933 when compared to the *E. coli* BCAT.

## Application to Biocatalysis

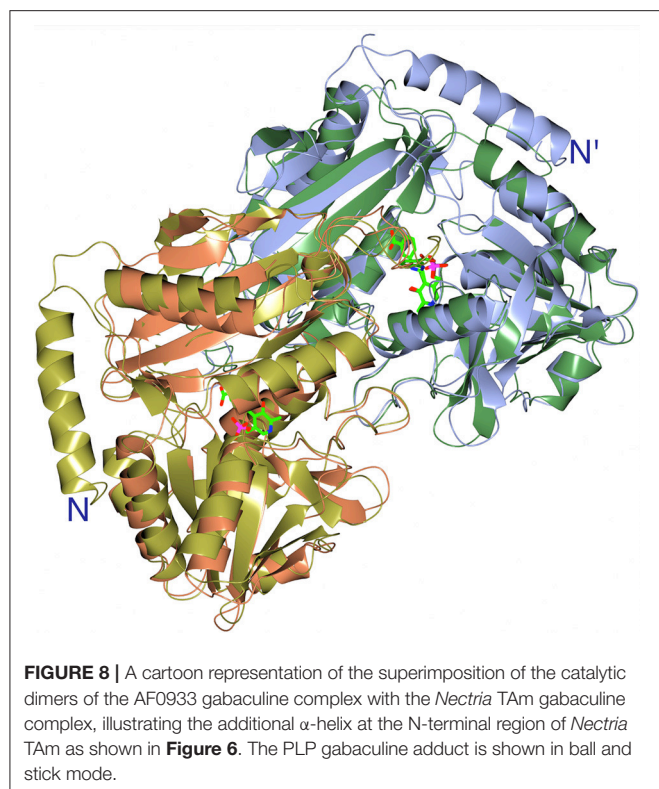
Both GEO1900 and AF0933 make a good thermostable and solvent tolerant scaffold which can be optimized for production of natural and synthetic chiral amines of interest to the pharmaceutical industries. The new BCAT enzymes have a chiral preference for the (*R*)-isomer of MBA but with low activity.

However, they offer the possibility to be further optimized for this and other substrates of interest using the knowledge of the active site cavity obtained from the current structural studies. The active site cavities of GEO1900 and AF0933 can be extended to better accommodate MBA and other bulky substrates. The comparison of GEO1900 and AF0933 with a (*R*)-specific TAM from a *Nectria* sp. shows that this fungal enzyme has an additional long N-terminal  $\alpha$ -helix that increases the depth of the active site of this enzyme. The presence of this N-terminal helix shown in **Figure 8**, a much tighter closure over the active site of the interdomain loop, and a different conformation of the loop between  $\beta 5$  and  $\beta 6$  of the N-terminal small domain, creates a difference in the substrate environment of the *Nectria* TAM in comparison with that in the GEO1900 and AF0933 BCATs. This would potentially alter the local dielectric constant within the fungal enzyme active site thereby facilitating the observed amination reactions of  $\omega$ -keto acids and different aldehydes and ketones by this TAM (Sayer et al., 2014).

## CONCLUSIONS

The approach developed in the THERMOGENE project was to search for these new archaeal TAM enzymes within the untapped resource of “natures” biodiversity in order to find enzymes that have evolved to be active under the extreme growth conditions of their host organisms. This approach has been successful and has resulted in the discovery of new thermostable BCATs displaying new features which would have been missed in a directed evolution approach using already identified mesophilic enzymes. The thermophilic archaea *G. acetivorans* and *A.*





*fulgidus* were isolated from different hydrothermal sources. The *G. acetivorans* has been isolated from samples of deep-sea hydrothermal vents of the Ashadze field, the deepest of the known World Ocean hydrothermal fields (Querellou et al., 2009) whereas the *A. fulgidus* was isolated in shallow hyperthermal sea water vents in Volcano Island in Sicily (Stetter, 1988) yet they have related BCAT enzymes with 79% sequence identity and similar properties. The GEO1900 and AF0933 enzymes have been characterized both biochemically and structurally. They both demonstrate significant thermostability and organic solvent stability and are promising candidates for industrial applications. Both the AF0933 and the GEO1900 are very thermostable as determined by the stability trials conducted. The AF0933 is a hexamer in solution as measured by gel filtration studies whereas the GEO1900 is mainly found as a dimer. The crystallographic studies have provided structures of both enzymes in the internal aldimine form, to 1.9 Å resolution for the *Geoglobus* enzyme and 2.0 Å for the *Archaeoglobus* enzyme. The structures of the enzymes with substrate and inhibitor have revealed the conformational changes occurring in the enzymes during catalytic turnover. A comparison of the *Archaeoglobus* inhibitor complex structure with the related (*R*)-selective fungal *Nectria* enzyme structure has produced further insights into differences in substrate specificity between the two enzymes. The new robust (*R*) selective BCAT enzymes will add to the enzyme “tool box” available for industrial biocatalysis and will have applications for production of new important chiral intermediates for the pharmaceutical industries using a sustainable route.

## MATERIALS AND METHODS

All reagents were purchased from Sigma–Aldrich, Buchs, Switzerland unless otherwise stated. The chromatography columns were obtained from GE Healthcare (Little Chalfont, UK).

### Expression and Purification

One gene encoding a class IV transaminase was identified in the complete genome sequence of *G. acetivorans* (Mardanov et al., 2011) by the BLAST search (Altschul et al., 1990). The nucleotide sequence of this gene was synthesized for optimal expression in *E. coli*. The GEO1900 was cloned into the *Bam*HI and *Hind*III sites of the expression vector pQE30 (Qiagen, Manchester, UK). The encoded recombinant protein contained an N-terminal MRGSHHHHHGS tag. The *E. coli* DLT1270 cells harboring the expression vector pQE30\_GEO1900 were grown in LB medium at 37°C, containing 100 µg ml<sup>−1</sup> ampicillin, to an OD<sub>600</sub> of 0.5. Protein expression was induced by addition of 1 mM Isopropyl-β-d-thiogalactopyranoside (IPTG) for 19 h at 37°C. The cells were harvested by centrifugation. The cell paste was resuspended in 50 mM Tris-HCl buffer, pH 7.5, containing 500 mM NaCl, 2 M urea, 20 mM imidazole, 2 mM β-mercaptoethanol, 0.1% (v/v) Triton X-100, and 1 mM phenylmethylsulfonyl fluoride (PMSF). After the cells were disrupted by sonication the crude cell extract was centrifuged at 10,000 g for 25 min. The filtered supernatant was applied to a Hi-Trap chelating HP column (2 ml) equilibrated in 50 mM Tris-HCl, pH 7.5, containing 500 mM NaCl, 20 mM imidazole, and 0.1% (v/v) Triton X-100. The (His)<sub>6</sub>-tagged recombinant GEO1900 was eluted with a linear gradient from 20 to 500 mM imidazole in the same buffer but without Triton X-100. The active fractions were pooled and incubated with 1 mM PLP for 1 h at 40°C, concentrated using a 30 kDa cut-off centrifugal filter device (Millipore), and applied to a Superdex 200 column (GE Healthcare) equilibrated in 30 mM Tris-HCl, pH 8.0, supplemented with 300 mM NaCl. The fractions containing the holo GEO1900 protein were stored at −20°C with the addition of 50% glycerol. The protein purity was analyzed by SDS-PAGE (12%). The protein concentration was determined spectrophotometrically using the calculated extinction coefficient of 24,955 M<sup>−1</sup> cm<sup>−1</sup>.

Another class IV transaminase gene was identified by searching known thermophilic archaeal genomes using the BCAT from *N. haematococca* as a query sequence. This identified a 290 amino acid protein in *A. fulgidus* (Klenk et al., 1997). The nucleotide sequence of this gene was synthesized for optimal codon usage for expression in *E. coli*. The expression of recombinant AF0933 was carried out using the plasmid pLATE51-AF0933 (Thermo Fisher Scientific, Ashford, UK) which was transformed into *E. coli* Rosetta (DE3)-pLysS (Novagene). Cells were grown in 1.6 l LB medium at 37°C to an OD<sub>600</sub> of 0.4–0.6 and expression was started by the addition of 1 mM IPTG. After 4.5 h of further growth, cells were harvested by centrifugation (max 9,800 × g at 18°C). The *E. coli* pellet was suspended in 50 mM Tris-HCl, pH 8.2, containing 300 mM NaCl and 10 mM imidazole (buffer A). Cell disruption was performed by passing through a French pressure cell (16,000



psi) followed by centrifugation (max  $48,000 \times g$  at  $4^{\circ}\text{C}$ ). The supernatant was incubated for 20 min at  $70^{\circ}\text{C}$  to precipitate *E. coli* proteins and centrifuged (max  $48,000 \times g$  at  $4^{\circ}\text{C}$ ). The resultant supernatant was applied to a 1 ml nickel-nitrilotriacetic acid (Ni-NTA) column (Qiagen, Manchester, UK) equilibrated in buffer A using the ÄKTA FPLC system. After a washing step, protein was eluted by stepwise increase of the imidazole concentration up to 500 mM. The fraction containing the highest enzyme activity was applied to a Superdex 200 HiLoad size exclusion column ( $1.6 \times 60$  cm) equilibrated with 50 mM Tris-HCl, pH 7.5, containing 500 mM NaCl and 0.05 mM PLP (GE Healthcare, Little Chalfont, UK). The protein purity was analyzed by SDS-PAGE (12%). Protein concentration was determined by using the Bradford method (Bradford, 1976) or spectrophotometrically for the purified protein.

## Enzyme Activity Assays

The overall transamination reaction of GEO1900 was assayed by a discontinuous method by measuring the rate of formation of L-glutamic acid from AKG in the presence of BCAAs in the direct reaction or the formation of AKG from L-glutamate acid in the presence of keto acids in the reverse reaction. In the direct reaction assay GEO1900 (5–30  $\mu\text{g}$ ) was pre-incubated in 1 ml of the reaction mixture, containing 50 mM phosphate buffer, pH 8.0, 100 mM NaCl, 60  $\mu\text{M}$  PLP, and 5 mM of the amino acid at  $65^{\circ}\text{C}$  for 3 min. The reaction was initiated by the addition of 5 mM AKG. Samples (100–300  $\mu\text{l}$ ) were taken at different time points and frozen to stop the reaction. Then the L-glutamic acid concentration was evaluated in the samples after dilution to 1 ml by spectrophotometry employing the glutamate dehydrogenase (GluDH) assay at  $25^{\circ}\text{C}$  in 50 mM Tris-HCl buffer, pH 9.0, supplemented with 1 mM NAD and 0.5 U GluDH. In the reverse reaction assay the enzyme (5–30  $\mu\text{g}$ ) was preincubated in the reaction mixture (1 ml) containing 50 mM phosphate buffer, pH 8.0, 100 mM NaCl, 60  $\mu\text{M}$  PLP, and 5 mM L-glutamic acid at  $65^{\circ}\text{C}$  for 3 min. The reaction was initiated by the addition of 5 mM ketoacid. Samples (100–300  $\mu\text{l}$ ) were taken at several time points and frozen to stop the reaction. The  $\alpha$ -ketoglutarate concentration was evaluated in the samples after dilution to 1 ml by spectrophotometry employing the GluDH assay at  $25^{\circ}\text{C}$  in 20 mM sodium phosphate buffer, pH 7.3, supplemented with 0.09 mM NADH, 30 mM ammonium chloride and 0.5 U GluDH. The specific activity of GEO1900 was calculated from the initial linear region of the progress curve of the L-glutamic acid/  $\alpha$ -ketoglutarate formation.

Steady-state kinetic parameters for pyruvate and AKG were determined from the substrate saturation curves for the overall reaction between the keto substrate and 5 mM L-Leu as a co-substrate. The saturation curves were analyzed using the Michaelis-Menten model. The GEO1900 activity toward (R) and (S)-MBA was measured spectrophotometrically at  $65^{\circ}\text{C}$  by the acetophenone assay (Schätzle et al., 2011; Boyko et al., 2016).

The AF0933 BCAT activity was measured as 3-methyl-2-oxobutyrates dependent formation of AKG from L-glutamate. One unit of enzyme activity corresponded to the conversion of 1  $\mu\text{mol}$  of substrate consumed or product formed per min. The standard assay mixture contained 0.1 M Tris-HCl, pH 8,

0.1 mM PLP, 5 mM 3-methyl-2-oxobutyrates, 10 mM L-glutamate and protein. Samples were incubated at  $70^{\circ}\text{C}$  up to 40 min, the reaction was stopped on ice and the amount of AKG formed was quantified by measuring NADH consumption in the reductive amination of AKG to L-glutamate at 340 nm. The quantification mixture (0.3 ml) was incubated at  $20^{\circ}\text{C}$  for 10 min and contained 0.1 M Tris-HCl, pH 7.5, 0.3 mM NADH, 40 mM  $\text{NH}_4\text{Cl}$ , 5–100  $\mu\text{l}$  of the samples and 0.32 U of L-glutamate dehydrogenase. Protein was determined by the Bradford method.

Kinetic constants for 3-methyl-2-oxobutyrates, 4-methyl-2-oxopentanoate and 3-methyl-2-oxopentanoate were determined with substrate concentrations up to 15 mM, 10 mM L-glutamate and 0.7–4  $\mu\text{g}$  protein using standard conditions. Kinetic constants for L-glutamate were determined with substrate concentrations up to 80 mM, 5 mM 3-methyl-2-oxobutanoic acid and 2.7–9.1  $\mu\text{g}$  protein using standard conditions. The substrate specificity was tested with 1 or 10 mM amino acceptors, 2.5  $\mu\text{g}$  protein using standard conditions.

The reverse reaction was measured as L-valine dependent decrease of AKG. The assay mixture contained 0.1 M Tris-HCl, pH 8.0, 0.1 mM PLP, 10 mM L-valine, 10 mM AKG and 5.3  $\mu\text{g}$  protein. Standard conditions were used for further incubation and quantification.

AF0933 activity was also measured as 3-methyl-2-oxobutyrates dependent formation of pyruvate from L-alanine. The assay mixture contained 0.1 M Tris-HCl, pH 8.0, 0.1 mM PLP, 5 mM 3-methyl-2-oxobutyrates, up to 20 mM L-alanine and 2.7  $\mu\text{g}$  protein. Samples were incubated at  $70^{\circ}\text{C}$  up to 20 min, the reaction was stopped on ice and the amount of pyruvate formed was quantified by measuring the reduction of pyruvate to lactate at 340 nm. The quantification mixture (0.3 ml) was incubated at  $20^{\circ}\text{C}$  for 10 min and contained 0.1 M Tris-HCl, pH 7.5, 0.3 mM NADH, 15–100  $\mu\text{l}$  of the samples and 1.8 U of L-lactate dehydrogenase (Roche, Burgess Hill, UK).

The AF0933 activity with (R) and (S)-MBA was determined by the acetophenone assay. The assay mixture of 1 ml contained 0.1 M Tris-HCl, pH 8.0, 0.1 mM PLP, 2 mM 3-methyl-2-oxobutyrates, 2-oxobutyrates or pyruvate as amino acceptor, 2 mM (R)-MBA or (S)-MBA as amino donor and 16  $\mu\text{g}$  protein. Samples were incubated at  $65^{\circ}\text{C}$  up to 180 min, the reaction was stopped on ice and the absorbance was measured at 245 nm. The molar extinction coefficient of  $11.6 \text{ mM}^{-1} \text{ cm}^{-1}$  was used for acetophenone.

## Effects of pH, Temperature, and Solvents on the Transamination Reaction

The pH optimum of the reaction for GEO1900 with 5 mM L-Leu, and 5 mM AKG was determined at  $65^{\circ}\text{C}$  using 50 mM sodium phosphate, pH 6.0–8.0, 50 mM BICINE containing 100 mM NaCl.

The temperature optimum of the transamination reaction between 5 mM L-Leu, and 5 mM AKG was determined in the temperature range from 40 to  $95^{\circ}\text{C}$  in 50 mM phosphate buffer, pH 7.9, supplemented with 100 mM NaCl. The thermostability of GEO1900 was determined by incubating a 1.0 mg/mL enzyme samples in 100 mM phosphate buffer, pH 8.0 at different

temperatures between 30 and 70°C for 2 h. After the thermal treatment, the remaining enzyme activity was determined at 65°C as described above.

The pH optimum of AF0933 was determined with 3.6 µg protein, standard conditions and the following buffers each at a concentration of 0.1 M: MES (pH 5.5, 6.0, 6.5), Bis-Tris (pH 6.5, 7.0), HEPES (pH 7.0, 7.5, 8.0), Tris-HCl (pH 7.5, 8.0, 8.5), Bicine (pH 8.5, 9.0) and Piperazine (pH 9.0, 9.5, 10.0).

The temperature optimum of the reaction of AF0933 was determined between 50°C and 95°C in 0.1 M HEPES, pH 7.0, containing 0.1 mM PLP, 5 mM 3-methyl-2-oxobutyrates, 10 mM L-glutamate and 2.7–5.4 µg protein.

The thermostability of AF0933 was measured between 55 and 95°C, therefore protein was incubated in 3 volumes of buffer (0.1 M HEPES, pH 7.0 adjusted to the respective temperature) up to 120 min. Samples (2.8 µg protein) were chilled on ice and remaining BCAT activity was determined in 0.1 M HEPES, pH 7.0 using standard conditions.

The influence of solvents (ethanol, DMSO, methanol and acetonitrile) on activity was analyzed between 10 and 50% solvent. The respective solvents were incubated with 2.2 µg protein (for AF0933) and 50 µg protein (for GEO1900) for 1 h at 20°C and 55°C then remaining activity was measured with standard conditions.

## Protein Crystallization

For crystallization fractions of GEO1900 after the size-exclusion chromatography step were concentrated up to 10 mg/ml using an Amicon Ultra 30 kDa MWCO (Merck Millipore, Darmstadt, Germany) centrifugal filter units and stored at –70°C in 30 mM Tris-HCl buffer, pH 8.0, supplemented with 300 mM NaCl and 100 µM PLP. High-throughput crystallization screening was setup using a robotic system (Rigaku Automation, USA) at the Resource Centers division of the NBICS Centre of the National Research Centre “Kurchatov Institute” (Boyko et al., 2013). The sitting-drop vapor-diffusion method was applied at 20°C in 96-well CrystalMation plates (Rigaku Automation, USA). The first round of the screen resulted in three in the following conditions: (1) 0.2 M TMAO, 0.1 M Tris-HCl, pH 8.5, 20% PEG 2000MME; (2) 30 mM citric acid, pH 3.0; 70 mM Bis-tris propane, 20% PEG 3350; (3) 0.2 M KCl; 0.1 M ammonium sulfate, pH 7.5; 20% 2-propanol, pH 7.5. Crystals grown under conditions (1) and (2) had rod-like shape with bad surface morphology and tended to form clusters. Under condition (3) numerous cube-like crystals were grown with the dimensions of 20 × 20 × 20 µm. The further optimization was made using the hanging-drop vapor-diffusion technique in 24-well VDX plates (Hampton Research, Aliso Viejo, California, USA) at 20°C by varying ammonium sulfate concentration, pH of the reservoir solution and addition of organic compounds. Crystals suitable for X-ray diffraction appeared in 3 months in the 3 µl drop with protein to precipitant ratio of 1:1 equilibrated over 500 µl of reservoir solution. The crystals used for data collection grew in (1) 0.1 M Tris-HCl pH 8.0, 12% glycerol, 1.5 M ammonium sulfate, and (2) 0.1 M HEPES, pH 7.5; 0.7 M lithium sulfate. Crystals of GEO1900 grown under condition (1) (GEO1900\_holo) were briefly soaked in mother liquor containing 30% glycerol as a

cryoprotectant. To obtain the complex of GEO1900 with AKG crystals from condition (2) were soaked for 20 min in the mother liquor supplemented with 50 mM AKG (GEO1900\_AKG) then transferred into a cryoprotectant containing the same solution with addition of 30% ethylene glycol. Crystals were flash-cooled to 100 K in liquid nitrogen immediately prior to the X-ray experiment.

Prior to crystallization the AF0933 enzyme was concentrated to ~10 mg ml<sup>–1</sup> using a 10 kDa Vivaspinn membrane (Vivaproducts, Littleton, Massachusetts, USA). Microbatch crystallization trials were set up using an Oryx 6 crystallization robot (Douglas Instruments, Hungerford, England) using the JCSG+ screen (Molecular Dimensions, Newmarket, England) (Newman et al., 2005). The droplet consisted of a 50:50 ratio of protein solution to screen solution and was covered with Al's oil (50:50 mixture of silicone oil and paraffin). The crystallization plates were stored at 18°C and were regularly checked for appearance of crystals under a light microscope.

Crystals of the AF0933 holoenzyme were grown using several different PEG and methylpentanediol (MPD) conditions. The best crystals grew in 100 mM cacodylate buffer pH 6.5, 200 mM magnesium chloride and 50% v/v PEG200. To obtain the inhibitor bound complex AF0933 was co-crystallized with 5 mM gabaculine under the same conditions.

## Data Collection and Processing

The GEO1900 data were collected at the BL41XU beamline of the Spring8 synchrotron (Harima, Japan) using a Pilatus detector. The data were indexed, integrated, and scaled using XDS (Kabsch, 2010) for the holo enzyme and using Mosflm (Battye et al., 2011) for the complex with AKG. The data showed both structures were not twinned, based on the L-test (Padilla and Yeates, 2003). The program Pointless (Evans, 2006) suggested the P3<sub>1</sub>21 space group or its enantiomorph P3<sub>2</sub>21.

Data for AF0933 were collected on beamline I04-1 at the Diamond Synchrotron light source (Oxford, UK) under cryo conditions (100 K in a stream of gaseous nitrogen). Data were processed using XDS (Kabsch, 2010) and AIMLESS (Evans and Murshudov, 2013) in the Xia2 pipeline (Winter et al., 2013). Further data and model manipulation was carried out using the CCP4 suite of programs (Winn et al., 2011). The data collection and processing statistics for both GEO1900 and AF0933 are summarized in Table 3.

## Structure Solution and Refinement

The structure of GEO1900\_AKG was solved at 2.2 Å resolution by the molecular replacement (MR) method using the BALBES MR pipeline (Long et al., 2007), the MR search was carried out in both enantiomorph space groups. The solution was found in P3<sub>2</sub>21 with the atomic model of BCAT from *T. thermophilus* (PDB ID: 2EII), which shares 42% of amino acid sequence identity with GEO1900, and was refined to an *R*-free of 37.7% without any model rebuilding, the BALBES estimation of probability of correct solution was 99%. The GEO1900\_holo structure was solved to 2.2 Å by MR with the atomic coordinates of GEO1900\_AKG.

The AF0933 holo-structure was phased by MR implemented in MOLREP (Vagin and Teplyakov, 2010) using a sequence modified model (Lebedev et al., 2008) of a dimer of *N. haematococca* Tam [31% sequence identity; PDB ID 4CMD; (Sayer et al., 2014)] as a model. The gabaculine inhibitor bound complex of AF0933 crystallized in the same space group  $P2_12_12_1$ .

Refinement of all structures was carried out with REFMAC5 (Murshudov et al., 2011). The manual rebuilding of the models into electron density maps was carried out with the COOT interactive graphics program (Emsley et al., 2010). The final refinement cycles of the GEO1900\_holo structure were conducted with inclusion of hydrogen atoms in riding positions. The 6-fold NCS averaging by DM (Cowtan, 2010) was used for phase improvement of AF0933 structures. The resulting phases from density modification were used for phase refinement in REFMAC5 (Pannu et al., 1998). BUSTER refinement maps (Bricogne et al., 2016) were used for building better positioning of poorly defined loops and of the inhibitor with partial occupancy in the AF0933 active site.

## Structure Analysis and Validation

PROCHECK was used for assessment of the quality of the models (Laskowski et al., 1993). Images were created with CCP4mg (McNicholas et al., 2011) and PyMOL molecular graphics system (Schrödinger, LLC). A visual inspection of the models was performed using the PyMOL and COOT software. The BLAST service (Altschul et al., 1990) was used for the amino acid alignment. The structure superposition was carried out using the PDBeFOLD program (Krissinel and Henrick, 2004). The contacts were analyzed using WHATIF (Vriend, 1990) and PISA (Krissinel and Henrick, 2007) software. The ligand coordination was visualized by LigPlot (Laskowski and Swindells, 2011).

The atomic coordinates and structure factors for both GEO1900 and AF0933 have been deposited in the Protein Data Bank. The PDB codes for GEO1900 holo enzyme as 5CM0 and GEO1900 complexed with AKG 5E25. The PDB codes for AF0933 holo enzyme 5MQZ and AF0933 complexed with gabaculine 5MR0.

## REFERENCES

- Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* 67, 271–281. doi: 10.1107/S0907444910048675
- Bezudnova, E. Y., Boyko, K. M., and Popov, V. O. (2017). Properties of bacterial and archaeal branched-chain amino acid aminotransferases. *Biochemistry* 13, 1572–1591. doi: 10.1134/S0006297917130028
- Boyko, K. M., Lipkin, A. V., Popov, V. O., and Kovalchuk, M. V. (2013). From gene to structure: the protein factory of the NBICS Centre of Kurchatov Institute. *Crystallogr. Rep.* 58, 442–449. doi: 10.1134/S106377451105004X
- Boyko, K. M., Stekhanova, T. N., Nikolaeva, A. Y., Mardanov, A. V., Rakitin, A. L., Ravin, N. V., et al. (2016). First structure of archaeal branched-chain amino acid aminotransferase from *Thermoproteus uzoniensis* specific for L-amino acids and R-amines. *Extremophiles* 20, 215–225. doi: 10.1007/s00792-016-0816-z
- Bradford, M. M. (1976). Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3
- Braunstein, A. E., and Shemyakin, M. M. (1953). The theory of processes of amino acid metabolism catalysed by pyridoxal 5'-phosphate-dependent enzymes. *Biokhimiya* 18, 393–411.
- Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, P., et al. (2016). *BUSTER Version 2.10.3*. Cambridge: Global Phasing Ltd.
- Chen, C. D., Lin, C. H., Chuankhayan, P., Huang, Y. C., Hsieh, Y. C., Huang, T. F., et al. (2012). Crystal structures of complexes of the branched-chain aminotransferase from *Deinococcus radiodurans* with alpha-ketoisocaproate and L-glutamate suggest the radiation resistance of this enzyme for catalysis. *J. Bacteriol.* 194, 6206–6216. doi: 10.1128/JB.01659-12
- Cho, B. K., Seo, J. H., Kang, T. W., and Kim, B. G. (2003). Asymmetric synthesis of L-homophenylalanine by equilibrium-shift using recombinant aromatic L-amino acid transaminase. *Biotechnol. Bioeng.* 83, 226–234. doi: 10.1002/bit.10661
- Cowtan, K. (2010). Recent developments in classical density modification. *Acta Crystallogr. D Biol. Crystallogr.* 66, 470–478. doi: 10.1107/S090744490903947X

## AUTHOR CONTRIBUTIONS

EB, NR, AM, PJ, and CS performed the discovery, cloning, and expression of the novel enzymes, while EB, PJ, CS, AN, and KB performed the protein purification, crystallization, and structure determination studies with MI. KB carrying out data analysis and structural refinement. AN, TS, MS, and J-MS performed the biochemical characterization of functional properties and the substrate specificity studies with PS. JL, VP, PS, MI, and EB coordinated the work and wrote the manuscript with input from all authors.

## FUNDING

This work was funded by the ERA-IB project THERMOGENE. The crystallographic study of AF0933 was carried out at Exeter, UK with support from the University of Kiel, Germany for the enzyme kinetic analysis. The Exeter JAL group would like to thank the BBSRC funded grant BB/L002035/1 and the University of Exeter for support.

## ACKNOWLEDGMENTS

The PS group would like to thank the BMBF grant 031A22 and the University of Kiel for support. The Exeter group would also like to thank the Diamond Synchrotron Light Source for access to beamline I04-1 (proposals No. MX8889) and the beamline scientists for assistance. The biochemical and crystallographic studies of GEO1900 were carried out in the Research Center of Biotechnology of the Russian Academy of Science in Moscow who would like to thank support from the Russian Federal Space Agency, and the Russian Science Foundation (project No 14-24-00172, in part of protein purification and structural work).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbioe.2019.00007/full#supplementary-material>



- Dementieva, I. S., Zakomirdina, L. N., Sinitzina, N. I., Antson, A. A., Wilson, K. S., Isupov, M. N., et al. (1994). Crystallization and preliminary X-ray investigation of holotryptophanases from *Escherichia coli* and *Proteus vulgaris*. *J. Mol. Biol.* 235, 783–786. doi: 10.1006/jmbi.1994.1033
- Diederichs, K., and Karplus, P. A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat. Struct. Biol.* 4, 269–275. doi: 10.1038/nsb0497-269
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501. doi: 10.1107/S0907444910007493
- Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* 62, 72–82. doi: 10.1107/S0907444905036693
- Evans, P. R., and Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* 69, 1204–1214. doi: 10.1107/S0907444913000061
- Goto, M., Miyahara, I., Hayashi, H., Kagamiyama, H., and Hirotsu, K. (2003). Crystal structures of branched-chain amino acid aminotransferase complexed with glutamate and glutarate: true reaction intermediate and double substrate recognition of the enzyme. *Biochemistry* 42, 3725–3733. doi: 10.1021/bi026722f
- Guan, L. J., Ohtsuka, J., Okai, M., Miyakawa, T., Mase, T., Zhi, Y., et al. (2015). A new target region for changing the substrate specificity of amine transaminases. *Sci. Rep.* 5:10753. doi: 10.1038/srep10753
- Hayashi, H. (1995). Pyridoxal enzymes mechanistic diversity and uniformity. *J. Biochem.* 118, 463–473. doi: 10.1093/oxfordjournals.jbchem.a124931
- Hirotsu, K., Goto, M., Okamoto, A., and Miyahara, I. (2005). Dual substrate recognition of aminotransferases. *Chem. Rec.* 5, 160–172. doi: 10.1002/tcr.20042
- Höhne, M., Schatzle, S., Jochens, H., Robins, K., and Bornscheuer, U. T. (2010). Rational assignment of key motifs for function guides *in silico* enzyme identification. *Nat. Chem. Biol.* 6, 807–813. doi: 10.1038/nchembio.447
- Humble, M. S., Cassimjee, K. E., Håkansson, M., Kimbung, Y. R., Walse, B., Abedi, V., et al. (2012). Crystal structures of the *Chromobacterium violaceum*  $\omega$ -transaminase reveal major structural rearrangements upon binding of coenzyme PLP. *FEBS J.* 279, 779–792. doi: 10.1111/j.1742-4658.2012.08468.x
- Hutchinson, E. G., and Thornton, J. M. (1996). PROMOTIF—a program to identify and analyze structure motifs in proteins. *Protein Sci.* 5, 212–220. doi: 10.1002/pro.5560050204
- Hutson, S. (2001). Structure and function of branched-chain aminotransferases. *Prog. Nucl. Acid Res. Mol. Biol.* 70, 175–206. doi: 10.1016/S0079-6603(01)70017-7
- Inoue, K., Kuramitsu, S., Aki, K., Watanabe, Y., Takagi, T., Nishigai, M., et al. (1988). Branched-chain amino acid aminotransferase of *Escherichia coli*: overproduction and properties. *J. Biochem.* 104, 777–784. doi: 10.1093/oxfordjournals.jbchem.a122549
- Jansonius, J. N. (1998). Structure evolution and action of vitamin B6-dependent enzymes. *Curr. Opin. Struct. Biol.* 8, 759–769. doi: 10.1016/S0959-440X(98)80096-1
- Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66, 125–132. doi: 10.1107/S0907444909047337
- Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637. doi: 10.1002/bip.360221211
- Kanda, M., Hori, K., Kurotsu, T., Ohgishi, K., Hanawa, T., and Saito, Y. (1995). Purification and properties of branched chain amino acid aminotransferase from gramicidin S-producing *Bacillus brevis*. *J. Nutr. Sci. Vitaminol.* 41, 51–60. doi: 10.3177/jnsv.41.51
- Karplus, P. A., and Diederichs, K. (2012). Linking crystallographic model and data quality. *Science* 336, 1030–1033. doi: 10.1126/science.1218231
- Kaulmann, U., Smithies, K., Smith, M. E. B., Hailes, H. C., and Ward, J. M. (2007). Substrate spectrum of  $\omega$ -transaminase from *Chromobacterium violaceum* DSM30191 and its potential for biocatalysis. *Enzyme Microb. Technol.* 41, 628–637. doi: 10.1016/j.enzmictec.2007.05.011
- Klenk, H. P., Clayton, R. A., Tomb, J. F., White, O., Nelson, K. E., Ketchum, K. A., et al. (1997). The complete genome sequence of the hyper-thermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* 390, 364–370. doi: 10.1038/37052
- Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2256–2268. doi: 10.1107/S0907444904026460
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797. doi: 10.1016/j.jmb.2007.05.022
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283–291. doi: 10.1107/S0021889892009944
- Laskowski, R. A., and Swindells, M. B. (2011). LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* 51, 2778–2786. doi: 10.1021/ci200227u
- Lebedev, A. A., Vagin, A. A., and Murshudov, G. N. (2008). Model preparation in MOLREP and examples of model improvement using X-ray data. *Acta Crystallogr. D Biol. Crystallogr.* 64, 33–39. doi: 10.1107/S0907444907049839
- Lee-Peng, F.-C., Hermodson, M. A., and Kohlhaw, G. B. (1979). Transaminase B from *Escherichia coli*: quaternary structure, amino-terminal sequence, substrate specificity, and absence of a separate valine- $\alpha$ -ketoglutarate activity. *J. Bacteriol.* 139, 339–345.
- Leuchtenberger, W., Huthmacher, K., and Drauz, K. (2005). Biotechnological production of amino acids and derivatives: current status and prospects. *Appl. Microbiol. Biotechnol.* 69, 1–8. doi: 10.1007/s00253-005-0155-y
- Littlechild, J., Novak, H., James, P., and Sayer, C. (2013). “Thermophilic microbes in environmental and industrial biotechnology,” in *Biotechnology of Thermophiles*, eds T. Satyanarayana, Y. Kawarabayasi, and J. Littlechild (Dordrecht: Springer), 481–509.
- Littlechild, J. A., Guy, J., Connelly, S., Mallett, L., Waddell, S., Rye, C. A., et al. (2007). Natural methods of protein stabilization: thermostable biocatalysts. *Biochem. Soc. Trans.* 35, 1558–1563. doi: 10.1042/BST0351558
- Long, F., Vagin, A. A., Young, P., and Murshudov, G. N. (2007). BALBES: a molecular-replacement pipeline. *Acta Crystallogr. D Biol. Crystallogr.* 64, 125–132. doi: 10.1107/S0907444907050172
- Lyskowski, A., Gruber, C., Steinkellner, G., Schürmann, M., and Schwab, H., et al. (2014). Crystal structure of an (R)-selective  $\omega$ -transaminase from *Aspergillus terreus*. *PLoS ONE* 9:e87350. doi: 10.1371/journal.pone.0087350
- Malik, M. S., Park, E. S., and Shin, J. S. (2012). Features and technical applications of  $\omega$ -transaminases. *Appl. Microbiol. Biotechnol.* 94, 1163–1171. doi: 10.1007/s00253-012-4103-3
- Mardanov, A. V., Gumerov, V. M., Beletsky, A. V., Prokofeva, M. I., Bonch-Osmolovskaya, E. A., Ravin, N. V., et al. (2011). Complete genome sequence of the thermoacidophilic crenarchaeon *Thermoproteus uzoniensis* 768-20. *J. Bacteriol.* 193, 3156–3157. doi: 10.1128/JB.00409-11
- McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. M. (2011). Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67, 386–394. doi: 10.1107/S0907444911007281
- Mehta, P. K., Hale, T. I., and Christen, P. (1993). Aminotransferases: demonstration of homology and division into evolutionary subgroups. *Eur. J. Biochem.* 214, 549–561. doi: 10.1111/j.1432-1033.1993.tb17953.x
- Metzler, D. E., Ikawa, M., and Snell, E. E. (1954). A general mechanism for vitamin B6-catalyzed reactions. *J. Am. Chem. Soc.* 76, 648–652. doi: 10.1021/ja01632a004
- Midelfort, K. S., Kumar, R., Han, S., Karmilowicz, M. J., McConnell, K., Gehlhaar, D. K., et al. (2013). Redesigning and characterizing the substrate specificity and activity of *Vibrio fluvialis* aminotransferase for the synthesis of imigabalin. *Protein Eng. Des. Sel.* 26, 25–33. doi: 10.1093/protein/gzs065
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., et al. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* 67, 355–367. doi: 10.1107/S0907444911001314
- Mutti, F. G., Fuchs, C. S., Pressnitz, D., Sattler, J. H., and Kroutil, W. (2011). Stereoselectivity of four (R)-selective transaminases for the asymmetric amination of ketones. *Adv. Synth. Catal.* 353, 3227–3233. doi: 10.1002/adsc.201100558
- Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I., Ben Jelloul, M., et al. (2005). Towards rationalization of crystallisation screening for small- to medium-sized laboratories: the PACT/JCSG+ strategy. *Acta Crystallogr. D Biol. Crystallogr.* 61, 1426–1431. doi: 10.1107/S0907444905024984



- Norton, J. E., and Sokatch, J. R. (1970). Purification and partial characterization of the branched chain amino acid transaminase of *Pseudomonas aeruginosa*. *Biochim. Biophys. Acta* 206, 261–269. doi: 10.1016/0005-2744(70)90109-9
- Okada, K., Hirotsu, K., Hayashi, H., and Kagamiyama, H. (2001). Structures of *Escherichia coli* branched-chain amino acid aminotransferase and its complexes with 4-methylvalerate and 2-methylleucine: induced fit and substrate recognition of the enzyme. *Biochemistry* 40, 7453–7463. doi: 10.1021/bi010384l
- Padilla, J. E., and Yeates, T. O. (2003). A statistic for local intensity differences: robustness to anisotropy and pseudo-centering and utility for detecting twinning. *Acta Crystallogr. D Biol. Crystallogr.* 59, 1124–1130. doi: 10.1107/S0907444903007947
- Pannu, N. S., Murshudov, G. N., Dodson, E. J., and Read, R. J. (1998). Incorporation of prior phase information strengthens maximum-likelihood structure refinement. *Acta Crystallogr. D Biol. Crystallogr.* 54, 1285–1294. doi: 10.1107/S0907444998004119
- Peisach, D., Chipman, D. M., Van Ophem, P. W., Manning, J. M., and Ringe, D. (1998). Crystallographic study of steps along the reaction pathway of D-amino acid aminotransferase. *Biochemistry* 37, 4958–4967. doi: 10.1021/bi972884d
- Pressnitz, D., Fuchs, C. S., Sattler, J. H., Knaus, T., Macheroux, P., Mutti, F. G., et al. (2013). Asymmetric amination of tetralone and chromanone derivatives employing  $\omega$ -transaminases. *ACS Catal.* 3, 555–559. doi: 10.1021/cs400002d
- Punta, M., Coghill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., et al. (2012). The Pfam protein families database. *Nucleic Acids Res.* 40, D290–D301. doi: 10.1093/nar/gkr1065
- Querellou, J., Bonch-Osmolovskaya, E. A., and Slobodkin, A. I. (2009). *Geoglobus acetivorans* sp. nov., an iron (III)-reducing archaeon from a deep-sea hydrothermal vent. *Int. J. Syst. Evol. Microbiol.* 59, 2880–2883. doi: 10.1099/ijs.0.011080-0
- Ramakrishnan, C., and Ramachandran, G. N. (1965). Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. *Biophys. J.* 5, 909–933. doi: 10.1016/S0006-3495(65)86759-5
- Richardson, J. S. (1981). The anatomy and taxonomy of protein structure. *Adv. Prot. Chem.* 34, 167–339. doi: 10.1016/S0065-3233(08)60520-3
- Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, 320–324. doi: 10.1093/nar/gku316
- Ruzicka, F. J., and Frey, P. A. (2010). Kinetic and spectroscopic evidence of negative cooperativity in the action of lysine 2,3-aminomutase. *J. Phys. Chem. B* 114, 16118–16124. doi: 10.1021/jp103856m
- Savile, C. K., Janey, J. M., Mundorff, E. C., Moore, J. C., Tam, S., Jarvis, W. R., et al. (2010). Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* 329, 305–309. doi: 10.1126/science.1188934
- Sayer, C., Bommer, M., Isupov, M., Ward, J., and Littlechild, J. (2012). Crystal structure and substrate specificity of the thermophilic serine:pyruvate aminotransferase from *Sulfolobus solfataricus*. *Acta Crystallogr. D Biol. Crystallogr.* 68, 763–772. doi: 10.1107/S0907444912011274
- Sayer, C., Isupov, M. N., Westlake, A., and Littlechild, J. A. (2013). Structural studies of *Pseudomonas* and *Chromobacterium*  $\omega$ -aminotransferases provide insights into their differing substrate specificity. *Acta Crystallogr. D Biol. Crystallogr.* 69, 564–576. doi: 10.1107/S0907444912051670
- Sayer, C., Martinez-Torres, R. J., Richter, N., Isupov, M. N., Hailes, H. C., Littlechild, J. A., et al. (2014). The substrate specificity, enantioselectivity and structure of the (R)-selective amine : pyruvate transaminase from *Nectria haematococca*. *FEBS J.* 281, 2240–2253. doi: 10.1111/febs.12778
- Schätzle, S., Steffen-Munsberg, F., Thontowi, A., Höhne, M., Robins, K., et al. (2011). Enzymatic asymmetric synthesis of enantiomerically pure aliphatic, aromatic and arylaliphatic amines with (R)-selective amine transaminases. *Adv. Synth. Catal.* 353, 2439–2445. doi: 10.1002/adsc.201100435
- Shin, J. S., Yun, H., Jang, J. W., Park, I., and Kim, B. G. (2003). Purification, characterization, and molecular cloning of a novel amine:pyruvate transaminase from *Vibrio fluvialis* JS17. *Appl. Microbiol. Biotechnol.* 61, 463–471. doi: 10.1007/s00253-003-1250-6
- Stetefeld, J., Jenny, M., and Burkhard, P. (2006). Intersubunit signaling in glutamate-1-semialdehyde-aminomutase. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13688–13693. doi: 10.1073/pnas.0600306103
- Stetter, K. O. (1988). *Archaeoglobus fulgidus* gen. nov., sp. nov. a new taxon of extremely thermophilic archaeobacteria. *Syst. Appl. Microbiol.* 10, 172–173. doi: 10.1016/S0723-2020(88)80032-8
- Tachiki, T., and Tochikura, T. (1973). Separation of L-leucine-pyruvate and L-leucine-alpha-ketoglutarate transaminases in *Acetobacter suboxydans* and identification of their reaction products. *Agric. Biol. Chem.* 37, 1439–1448. doi: 10.1080/00021369.1973.10860842
- Taylor, P. P., Pantaleone, D. P., Senkpeil, R. F., and Fotheringham, I. G. (1998). Novel biosynthetic approaches to the production of unnatural amino acids using transaminases. *Trends Biotechnol.* 16, 412–418. doi: 10.1016/S0167-7799(98)01240-2
- Thomsen, M., Skalden, L., Palm, G. J., Höhne, M., Bornscheuer, U. T., and Hinrichs, W. (2014). Crystallographic characterization of the (R)-selective amine transaminase from *Aspergillus fumigatus*. *Acta Crystallogr. D Biol. Crystallogr.* 70, 1086–1093. doi: 10.1107/S1399004714001084
- Tufvesson, P., Lima-Ramos, J., Jensen, J. S., Al-Haque, N., Neto, W., and Woodley, J. M. (2011). Process considerations for the asymmetric synthesis of chiral amines using transaminases. *Biotechnol. Bioeng.* 108, 1479–1493. doi: 10.1002/bit.23154
- Uchida, Y., Hayashi, H., Washio, T., Yamasaki, R., Kato, S., and Oikawa, T. (2014). Cloning and characterization of a novel fold-type I branched-chain amino acid aminotransferase from the hyper-thermophilic archaeon *Thermococcus* sp. *CKU-1 Extremophiles* 18, 589–602. doi: 10.1007/s00792-014-0642-0
- Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* 66, 22–25. doi: 10.1107/S0907444909042589
- Vaguine, A. A., Richelle, J., and Wodak, S. J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D Biol. Crystallogr.* 55, 191–205. doi: 10.1107/S0907444998006684
- van Oosterwijk, N., Willies, S., Hekelaar, J., Terwisscha van Scheltinga, A. C., Turner, N. J., and Dijkstra, B. W. (2016). Structural basis of the substrate range and enantioselectivity of two (S)-selective  $\omega$ -transaminases. *Biochemistry* 55, 4422–4431. doi: 10.1021/acs.biochem.6b00370
- Vriend, G. (1990). WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.* 8, 52–56. doi: 10.1016/0263-7855(90)80070-V
- Ward, J., and Wohlgemuth, R. (2010). High-yield biocatalytic amination reactions in organic synthesis. *Curr. Org. Chem.* 14, 1914–1927. doi: 10.2174/138527210792927546
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* 67, 235–242. doi: 10.1107/S0907444910045749
- Winter, G., Lobley, C. M. C., and Prince, S. M. (2013). Decision making in xia2. *Acta Crystallogr. D Biol. Crystallogr.* 69, 1260–1273. doi: 10.1107/S0907444913015308
- Yennawar, N., Dunbar, J., Conway, M., Hutson, S., and Farber, G. (2001). The structure of human mitochondrial branched-chain aminotransferase. *Acta Crystallogr. D Biol. Crystallogr.* 57, 506–515. doi: 10.1107/S090744490101925
- Yu, X., Wang, X., and Engel, P. C. (2014). The specificity and kinetic mechanism of branched-chain amino acid aminotransferase from *Escherichia coli* studied with a new improved coupled assay procedure and the enzyme's potential for biocatalysis. *FEBS J.* 281, 391–400. doi: 10.1111/febs.12609

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Isupov, Boyko, Sutter, James, Sayer, Schmidt, Schönheit, Nikolaeva, Stekhanova, Mardanov, Ravin, Bezudnova, Popov and Littlechild. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.